

A Role-Free Approach to Indexing Large RDF Data Sets in Secondary Memory for Efficient SPARQL Evaluation

George H. L. Fletcher and Peter W. Beck
 School of Engineering and Computer Science
 Washington State University, Vancouver, USA
 {fletcher, pwbeck}@wsu.edu

1 Introduction

Massive RDF data sets are becoming commonplace. RDF data is typically generated in social semantic domains (such as personal information management [2, 11, 13]) wherein a fixed schema is often not available a priori. We propose a simple *Three-way Triple Tree* (TripleT) secondary-memory indexing technique to facilitate efficient SPARQL query evaluation on such data sets. The novelty of TripleT is that (1) the index is built over the atoms occurring in the data set, rather than at a coarser granularity, such as whole triples occurring in the data set; and (2) the atoms are indexed regardless of the roles (i.e., subjects, predicates, or objects) they play in the triples of the data set. We show through extensive empirical evaluation that TripleT exhibits multiple orders of magnitude improvement over the state of the art on RDF indexing, in terms of both storage and query processing costs.

Preliminary Notions. We assume familiarity with the RDF and SPARQL standards [8, 12, 15], the B+tree data structure [4, 16], and the basics of conjunctive query processing [3, 16, 18]. Let \mathcal{A} be an enumerable set of atoms (e.g., Unicode strings). A *triple* is an element of $\mathcal{A} \times \mathcal{A} \times \mathcal{A}$. An RDF *graph* is a finite set of triples. For graph G , let

$$\begin{aligned}\mathcal{S}(G) &= \{s \mid (s, p, o) \in G\} \\ \mathcal{P}(G) &= \{p \mid (s, p, o) \in G\} \\ \mathcal{O}(G) &= \{o \mid (s, p, o) \in G\} \\ \mathcal{A}(G) &= \mathcal{S}(G) \cup \mathcal{P}(G) \cup \mathcal{O}(G).\end{aligned}$$

```

{⟨Yamada, authored, doc1⟩,
 ⟨Yamada, knows, McShea⟩,
 ⟨knows, is a kind of, social action⟩,
 ⟨Herzog, authored, doc2⟩,
 ⟨Herzog, authored, doc3⟩,
 ⟨McShea, performed, doc3⟩,
 ⟨McShea, past action, authored⟩,
 ⟨doc1, type, PDF⟩,
 ⟨doc1, rating, 4/5⟩,
 ⟨doc2, type, MP3⟩,
 ⟨doc3, type, MP3⟩,
 ⟨doc3, created on, 26.10.08⟩}

```

Figure 1: A triple graph.

The atoms appearing in $\mathcal{S}(G)$ are called the *subjects* of G ; the atoms appearing in $\mathcal{P}(G)$ are called the *predicates* of G ; and, the atoms appearing in $\mathcal{O}(G)$ are called the *objects* of G .

2 The Problem

The problem we consider in this paper is how to index a graph G to support efficient evaluation of *basic graph patterns* (BGP) over G . BGPs, which are conjunctions of *simple access patterns* (SAP), form the heart of all SPARQL queries.

Example 1 Consider the query “What are the dates and types of documents on which McShea was a performer?” over the triple store given in Figure 1. In SPARQL, where variables are identified by a leading $?$, this query can be formulated as follows:

```

SELECT ?date ?type
WHERE { McShea performed ?doc .
        ?doc created_on ?date .
        ?doc type ?type }

```

The WHERE clause of a SPARQL query specifies a BGP, which in this case consists of the conjunction of the following three SAPs:

(McShea, performed, ?doc), (?doc, created on, ?date), (?doc, type, ?type).

Conceptually, the evaluation of a BGP on a graph G consists of finding all variable bindings such that each of the BGP’s constituent SAPs simultaneously holds in G . In our example, there is only one set of valid variable bindings:

?doc	?date	?type
doc1	26.10.08	MP3

The **SELECT** clause indicates that only the bindings for ?date and ?type are returned in the query result.

The reader will recognize that BGPs are essentially conjunctive queries evaluated over a single ternary relation [3, 7, 9, 18, 21]. Joins between the SAPs of a BGP are induced by the co-occurrence of variables and atoms. There are six native BGP join types: subject-subject, subject-predicate, subject-object, predicate-predicate, predicate-object, and object-object joins. In Example 1, there is a subject-object join between the first SAP and both the second and third SAPs, due to the co-occurrence of variable ?doc. Furthermore, there is a subject-subject join between the second and third SAPs.

We specifically focus on the problem of designing *native* RDF index data structures to accelerate BGP evaluation. By native, we mean data structures which support the full range of BGP join patterns.

3 The Solution

Let G be a fixed RDF graph. In what follows, we use the B+tree secondary-memory data structure [4] to implement the various indexing techniques considered. However, any of a variety of appropriate secondary-memory data structures (e.g., linear hashing [16]) could also be also have been used.

3.1 State of the Art

To the best of our knowledge, the two major competitive proposals for native RDF indexing are multiple access patterns (MAP) and HexTree.

- **MAP.** In this approach, all three positions of triples are indexed: subjects (S), predicates (P), and objects (O), for some permutation of S, P, and O. MAP requires up to six separate indexes, corresponding to the six possible orderings of roles: SPO, SOP, PSO, POS, OSP, OPS. For example, for each $(s, p, o) \in G$, it is the case that $o\#p\#s$ is a

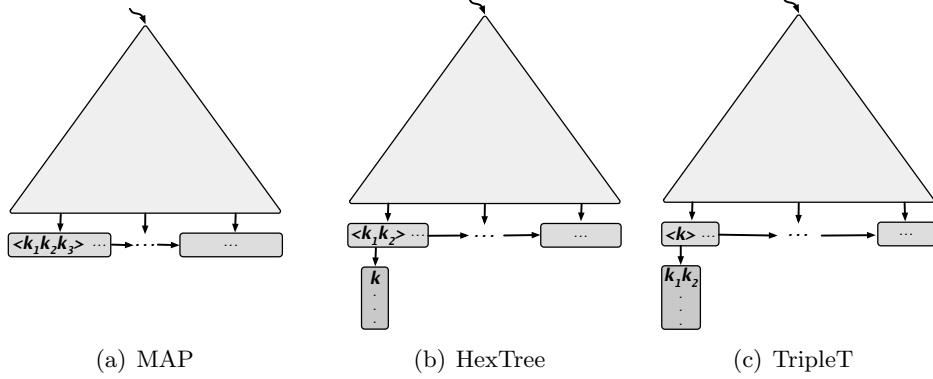


Figure 2: Varieties of Triple Trees.

key in the OPS index on G ; see Figure 2(a).¹ A BGP join evaluation requires two or more look-ups, potentially in different trees, followed by merge-joins. Major systems employing this technique include Virtuoso, YARS, RDF-3X, Kowari, and System-II [6, 10, 14, 26, 27]. In the present investigation we use the B+tree data structure for each of the MAP indexes (Figure 2(a)).

- **HexTree.** Recently in the Hexstore system, Weiss et al. [24] have proposed indexing two roles at a time. This approach requires up to six separate indexes corresponding to the six possible orderings of roles: SO, OS, SP, PS, OP, PO. Payloads are shared between indexes with symmetric orderings. For example, for each $(s, p, o) \in G$, it is the case that $s\#p$ is a key in the SP index on G , $p\#s$ is a key in the PS index on G , and both of these keys point to a payload of $\{o \in \mathcal{O}(G) \mid (s, p, o) \in G\}$; see Figure 2(b). As with MAP, join evaluation requires two or more look-ups, potentially in different trees, followed by merge-joins. Hexstore has only been proposed and evaluated as a main-memory data structure [24]. We propose HexTree as an effective secondary-memory realization of the Hexstore proposal using the B+tree data structure (Figure 2(b)).

Note that techniques have also been developed for indexing heuristically-selected classes of larger graph patterns, e.g., [23]. Such techniques, however, do not support processing of the full range of native BGP join patterns.

¹Where “#” is some reserved separator symbol.

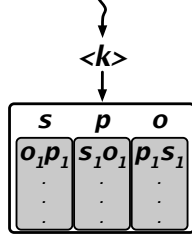


Figure 3: TripleT payload for atom k .

3.2 Our Proposal

We propose indexing the key-space $\mathcal{A}(G)$, regardless of the particular roles the atoms of $\mathcal{A}(G)$ play in the triples of G . For a key k , the payload is all triples of G in which atom k occurs (see Figure 2(c)). In particular, the payload for k consists of three “buckets”: one for all pairs (p, o) where $(k, p, o) \in G$, one for all pairs (s, o) where $(s, k, o) \in G$, and one for all pairs (s, p) where $(s, p, k) \in G$, (see Figure 3). In other words, there is one bucket apiece for all those triples where k occurs as a subject, for all those triples where k occurs as a predicate, and for all those triples where k appears as an object. For example, on the graph of Figure 1, the payload for `doc1` would consist of an object bucket $\langle (\text{Yamada}, \text{authored}) \rangle$, a subject bucket $\langle (4/5, \text{rating}), (\text{PDF}, \text{type}) \rangle$, and a predicate bucket $\langle \rangle$.² TripleT requires just one index, while efficiently supporting all join patterns native to SPARQL. For example, a subject-object join induced by the co-occurrence of an atom k can be evaluated by a single look-up on k followed by a merge-join between the subject and object buckets of k ’s payload. A join induced by the co-occurrence of a variable is implemented as multiple look-ups followed by merge-joins, as with MAP and HexTree. However, since the keys in TripleT are 1/3 the length of those in MAP and 1/2 those in HexTree, there is a significant increase in the branching factor of the TripleT B+tree, which leads to a significant reduction in cost for these look-ups.

TripleT does not favor any particular join types, supporting the full range of join patterns native to RDF data. The recently proposed “vertical-partitioning” approach [1] can be viewed as a special restricted case of TripleT where (1) only the atoms of $\mathcal{P}(G)$ are indexed and (2) only the predicate payload bucket for each key is maintained. In this sense, vertical-

²To facilitate query processing, note that we keep the pairs in each of the buckets sorted. By default, the subject bucket is sorted in OP order, the predicate bucket in SO order, and the object bucket in SP order.

partitioning is not a fully native RDF indexing technique; indeed, recent research has demonstrated practical limitations of this approach [19, 17, 24]. This research has also demonstrated similar limitations of the related “property table” RDF storage techniques [5, 20, 22, 25].

4 Empirical Evaluation

We implemented all three approaches using 8K blocks and 32-bit references, in virtual memory, using Python 2.5.2. All experiments were executed on a pair of 2.66 GHz dual-core Intel Xeon processors with 16 GB RAM running Mac OS X 10.4.11. Each experiment was performed using (1) simple synthetic data; (2) the DBPedia RDF data set; and, (3) the Uniprot RDF data set. Further details of these data sets are provided in the Appendix.

As mentioned above, in TripleT we only materialized the OP, SO, and SP sort orderings for the subject, predicate, and object payload buckets, respectively.³ Consequently, we only built the corresponding SOP, PSO, and OSP trees for MAP and the SO, PS, and OS trees for HexTree. In all of our experiments, the TripleT payloads occupied on average only one disk block. Hence, if a symmetric sort ordering was necessary for a merge join (e.g., if the PO ordering was necessary for the subject bucket while using TripleT or if the SPO ordering was necessary while doing a lookup in MAP), the sort was performed in main-memory without penalty.

4.1 Index size

In increments of 1 million triples, from 1 to 6 million triples, we built the three index types. The plots of the index sizes, in 8K blocks, are shown in Figures 4(a)-4(c). TripleT was up to eight orders of magnitude smaller, with a typical two orders of magnitude savings in storage cost. The reason for this can be attributed to (1) TripleT uses just one B+tree, whereas MAP and HexTree both require three B+trees, and (2) the key size in TripleT is 1/3 that of MAP and 1/2 that of HexTree, leading to significantly higher branching factor of the B+tree (and hence shallower trees).

³If necessary, each of the two possible sort orderings for each of the three TripleT buckets could be materialized. In this case, we would of course still need just one B+tree to index payloads.

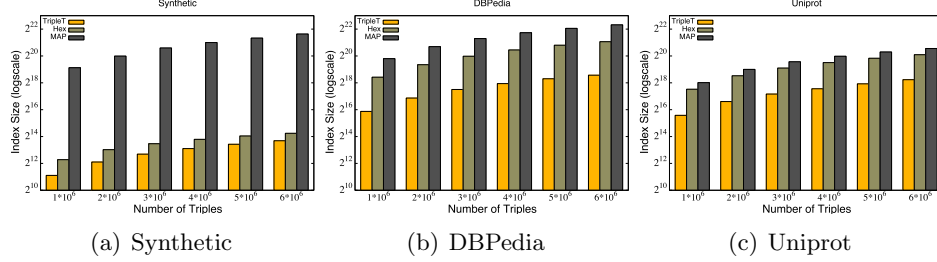


Figure 4: Index sizes, in 8K blocks.

4.2 Query performance

We use the classic I/O cost model for query evaluation, i.e., we use the number of block reads as our performance metric [16], as we are interested in comparing the technology-independent behavior of MAP, HexTree, and TripleT. We considered two query scenarios:

- A single SAP without variables, which we denote as a “ $k = 0$ ” join scenario. For each dataset, and for each size, we randomly selected ten triples from the dataset and recorded the costs of looking them up in MAP, HexTree, and TripleT. The average I/O cost of performing these lookups is given in Figures 5(a)-5(c).
- Basic BGP join patterns, which we denote as a “ $k = 1$ ” join scenario. We considered four sub-scenarios, covering the basic ways in which SAPs may be joined.
 1. Computing the join of two variable-free SAPs having one atom in common.
 2. Computing the join of two SAPs having one atom in common, one SAP having a single variable and the other variable-free.
 3. Computing the join of two SAPs having no atoms in common, each having a single variable, which they share.
 4. Computing the join of two SAPs having one atom in common, each having one variable, which they also share.

For each data set, for each size, we generated ten random BGPs of each of these four scenarios and recorded the cost of their evaluation

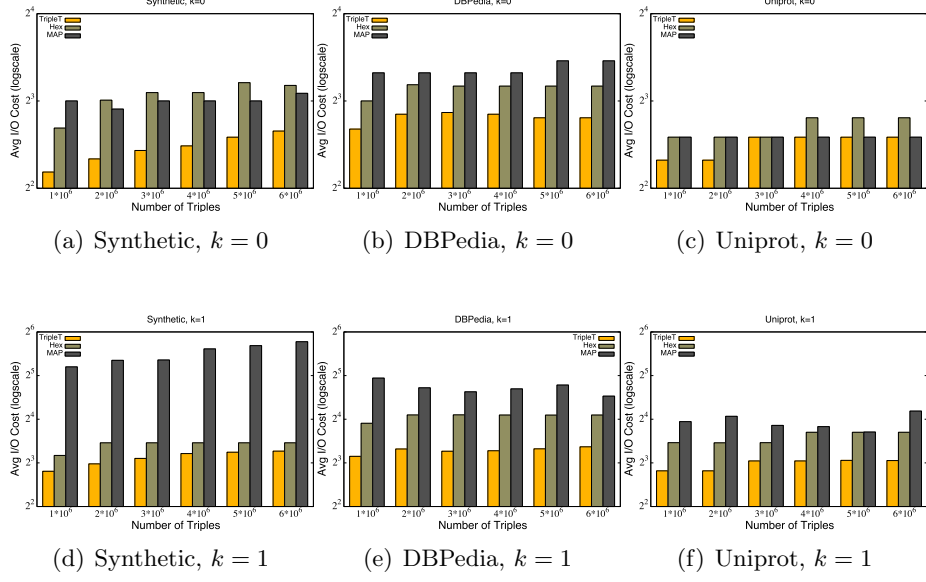


Figure 5: Cost of Query Processing.

using MAP, HexTree, and TripleT. The average I/O costs are given in Figures 5(d)-5(f).

We observe from these experiments that (1) for $k = 0$ TripleT never performed worse than MAP or HexTree, and usually better; and, (2) for $k = 1$, TripleT always out-performed MAP and HexTree, with up to two orders of magnitude improvement in I/O costs.

5 Concluding remarks

It is clear from this extensive evaluation of the full range of BGP join scenarios on both synthetic and real-world data sets that TripleT is a serious contender for indexing massive RDF data stores in secondary memory. Our proposal is conceptually quite simple, and hence straight forward to implement. Furthermore, TripleT exhibits multiple orders of magnitude improvement over the state of the art for both storage cost and query evaluation cost. In closing, we note that the many optimizations (such as various key compression schemes) which have been used in implementations of MAP and HexTree reported in the literature can equally be applied to TripleT.

References

- [1] Daniel J. Abadi, Adam Marcus, Samuel Madden, and Katherine J. Hollenbach. Scalable Semantic Web Data Management Using Vertical Partitioning. In *VLDB*, pages 411–422, Vienna, 2007.
- [2] Karl Aberer. Data Management in the Social Web. In *EDBT*, pages 1203–1204, Munich, 2006.
- [3] Ashok K. Chandra and Philip M. Merlin. Optimal Implementation of Conjunctive Queries in Relational Data Bases. In *ACM STOC*, pages 77–90, Boulder, CO, USA, 1977.
- [4] Douglas Comer. The Ubiquitous B-Tree. *ACM Comput. Surv.*, 11(2):121–137, 1979.
- [5] María del Mar Roldán García and José Francisco Aldana Montes. A Survey on Disk Oriented Querying and Reasoning on the Semantic Web. In *IEEE ICDE Workshop SWDB*, Atlanta, 2006.
- [6] Orri Erling. Towards Web Scale RDF. In *SSWS*, Karlsruhe, Germany, 2008.
- [7] George H. L. Fletcher. An Algebra for Basic Graph Patterns. In *LID*, Rome, 2008.
- [8] Tim Furche, Benedikt Linse, François Bry, Dimitris Plexousakis, and Georg Gottlob. RDF Querying: Language Constructs and Evaluation Methods Compared. In *Reasoning Web*, pages 1–52, Lisbon, Portugal, 2006.
- [9] Claudio Gutiérrez, Carlos A. Hurtado, and Alberto O. Mendelzon. Foundations of Semantic Web Databases. In *ACM PODS*, pages 95–106, Paris, 2004.
- [10] Andreas Harth, Jürgen Umbrich, Aidan Hogan, and Stefan Decker. YARS2: A Federated Repository for Querying Graph Structured Data from the Web. In *ISWC*, Busan, Korea, 2007.
- [11] David R. Karger, Karun Bakshi, David Huynh, Dennis Quan, and Vineet Sinha. Haystack: A General-Purpose Information Management Tool for End Users Based on Semistructured Data. In *CIDR*, pages 13–26, 2005.

- [12] Graham Klyne and Jeremy J. Carroll. Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation, 2004.
- [13] m c schraefel. What is an Analogue for the Semantic Web and Why is Having One Important? In *ACM Hypertext*, pages 123–132, Manchester, UK, 2007.
- [14] Thomas Neumann and Gerhard Weikum. RDF-3X: A RISC-Style Engine for RDF. In *VLDB*, Auckland, New Zealand, 2008.
- [15] Eric Prud’hommeaux and Andy Seaborne. SPARQL Query Language for RDF. W3C Recommendation, 2008.
- [16] Raghu Ramakrishnan and Johannes Gehrke. *Database Management Systems, 3rd Ed.* McGraw Hill, Boston, 2003.
- [17] Michael Schmidt, Thomas Hornung, Norbert Kuchlin, Georg Lausen, and Christoph Pinkel. An Experimental Comparison of RDF Data Management Approaches in a SPARQL Benchmark Scenario. In *ISWC*, pages 82–97, Karlsruhe, Germany, 2008.
- [18] Patricia G. Selinger, Morton M. Astrahan, Donald D. Chamberlin, Raymond A. Lorie, and Thomas G. Price. Access Path Selection in a Relational Database Management System. In *ACM SIGMOD*, pages 23–34, Boston, 1979.
- [19] Lefteris Sidiropoulos, Romulo Goncalves, Martin Kersten, Niels Nes, and Stefan Manigold. Column-Store Support for RDF Data Management: Not All Swans are White. In *VLDB*, Auckland, New Zealand, 2008.
- [20] Michael Sintek and Malte Kiesel. RDFBroker: A Signature-Based High-Performance RDF Store. In *ESWC*, pages 363–377, Budva, Montenegro, 2006.
- [21] Markus Stocker, Andy Seaborne, Abraham Bernstein, Christoph Kiefer, and Dave Reynolds. SPARQL Basic Graph Pattern Optimization Using Selectivity Estimation. In *ACM WWW*, pages 595–604, Beijing, 2008.
- [22] Yannis Theoharis, Vassilis Christophides, and Gregory Karvounarakis. Benchmarking Database Representations of RDF/S Stores. In *ISWC*, pages 685–701, Galway, Ireland, 2005.

- [23] Octavian Udrea, Andrea Pugliese, and V. S. Subrahmanian. GRIN: A Graph Based RDF Index. In *AAAI*, pages 1465–1470, Vancouver, B.C., 2007.
- [24] Cathrin Weiss, Panagiotis Karras, and Abraham Bernstein. Hexastore: Sextuple Indexing for Semantic Web Data Management. In *VLDB*, Auckland, New Zealand, 2008.
- [25] Kevin Wilkinson. Jena Property Table Implementation. In *SSWS*, pages 35–46, Athens, Georgia, USA, 2006.
- [26] David Wood, Paul Gearon, and Tom Adams. Kowari: A Platform for Semantic Web Storage and Analysis. In *XTech*, Amsterdam, 2005.
- [27] Gang Wu, Juanzi Li, and Kehong Wang. System II: a Hypergraph Based Native RDF Repository. In *WWW*, pages 1035–1036, Beijing, 2008.

Appendix

In this section we provide details of the data sets used in the experiments discussed in Section 4: (1) synthetic data, (2) the DBPedia RDF data set;⁴ and (3) the Uniprot RDF data set.⁵

For (1), we built two synthetic data sets of size 6 million (the results of Section 4 are the averages over these two sets). In the first set, we randomly generated n triples over $n^{1/3}$ unique atoms, for $n = 1,000,000$, to $n = 6,000,000$, in increments of one million, where repetitions of atoms were allowed within triples. In the second set, we randomly generated n triples over $\text{ceiling}(n^{1/3}) + 2$ unique atoms, for $n = 1,000,000$, to $n = 6,000,000$, in increments of one million, where repetitions of atoms within triples were disallowed.

For (2) and (3), we took an arbitrary sample of 10,000,000 triples from each data collection (treating the DBPedia infobox and pagelinks as one collection) — see Table 1. After cleaning and duplicate elimination, we kept 6,000,000 triples in each collection. In this cleaned data, we use only the first 400 (DBPedia) or 150 (Uniprot) characters of atoms (note that these are the basis of the fixed key sizes for the B+trees we built). This truncation only affected a few extremely long atoms appearing exclusively in the object position. Final statistics for these data sets are given in Table 2.

⁴<http://wiki.dbpedia.org>

⁵<http://dev.isb-sib.ch/projects/uniprot-rdf>

G	$ G $	average atom length
DBPedia	82,701,339	34.2
Uniprot	956,915,180	29.0

Table 1: Data sets

G	$ \mathcal{S}(G) $	$ \mathcal{P}(G) $	$ \mathcal{O}(G) $	$ \mathcal{A}(G) $	$ \mathcal{S}(G) \cap \mathcal{O}(G) $	$ \mathcal{S}(G) \cap \mathcal{P}(G) $	$ \mathcal{P}(G) \cap \mathcal{O}(G) $
DBPedia	1,370,679	20,873	1,848,114	2,852,484	387,182	0	0
Uniprot	4,357,005	81	1,734,176	5,644,939	446,311	0	12

Table 2: Basic statistics of sampled data sets